DOCUMENT RESUME

ED 143 621 SP 011 281

AUTHOR Hayes-Roth, Frederick
TITLE Learning By Example.
INSTITUTION Rand Corp., Santa Monica, Calif.
REPORT NO Rand-P-5835
PUB DATE Mar 77
NOTE 21p.; Not available in hard copy due to marginal legibility of the original document
AVAILABLE FROM Publications Department, The Rand Corporation, 1700 Main Street, Santa Monica, California 90406 ($1.50)

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Abstraction Levels; Abstract Reasoning; *Cognitive Processes; *Comprehension; *Concept Formation; Educational Research; *Learning Processes; Learning Theories; Psychological Patterns; Serial Learning; Thought Processes

ABSTRACT

One of the most typical ways in which people learn is by inferring general rules from examples. In recent years, significant progress has been made toward understanding how learning from examples can occur, determining when it does occur, and identifying conditions that promote it. This paper reviews these results and then suggests a program of research and application that exploits improved understanding of this powerful basis for learning. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED143621

SP

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

T. Cockrell

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

LEARNING BY EXAMPLE

Frederick Hayes-Roth

BEST COPY AVAILABLE

March 1977

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

P-5835

SP 011 281

Copyright © 1977
THE RAND CORPORATION

The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation
Santa Monica, California 90406

1. WHAT IS LEARNING BY EXAMPLE?

Everyone has many personal experiences of learning by example. While much psychological research has investigated "concept learning" (cf. Bruner, Goodnow, Austin, 1956; Hayes-Roth & Hayes-Roth, 1977; Hunt, 1952), that rubric is too narrow to embrace the variety of situations in which learning by example occurs. A brief list of such situations includes:

(1) Traditional concept learning, such as inducing the class characteristic of "triangle": "Three distinct line segments such that each line segment is co-terminus with a different line segment at each of its endpoints." Such a rule can be induced from various examples of triangles; all examples necessarily manifest the rule, although they may differ from one another in irrelevant ways (e.g., in absolute and relative size, shape, orientation, color, texture). Note that although most traditional concept learning tasks employ only attribute-value descriptions, this learning task requires higher-order, relational logic to characterize the structural constraints among the lines of a triangle.

(2) Serial pattern learning, such as predicting the next item in a conceptually organized sequence. Traditional research on this problem has centered on mathematical sequences of symbols and various algorithmic models of memory processes for simulating the sequence generator. Other examples of this type of behavior include anticipation of expectable events (e.g., words or topics

in a text that are predictable from preceding context) and prediction of cyclic phenomena. In these situations, subsequences of the preceding sequence of items serve as examples from which the sequence-generation rule is induced.

(3) Transformation learning, such as inducing the rule for converting active sentences to semantically equivalent passive sentences. Presumably, this type of general syntactic procedure is abstracted from examples of semantically equivalent pairs of sentences illustrating such a transformation. However, the use of transformations extends far beyond the domain of syntax. Newell and Simon have shown that similar types of rules (called productions) provide an excellent model of human problem solving activity. Their theory postulates that the problem solver applies these [condition->action] rules to internal representations of the problem whenever the properties of the problem state satisfy the rules' conditions. These rules too are presumably inferred by generalizing previous experiences, including both trial-and-error successes and instructive examples.

(4) Category learning, such as inducing the sets of words that are adjectives and those that are nouns and the associated rule of English grammar that adjectives may precede nouns but not vice versa. While such formal categories and grammatical principles may not be consciously understood, behavioral evidence of this sort of learning is ubiquitous. It is

hypothesized in this case too that learning results from experience of examples.

Learning by example, as a theory of the way concepts and general rules are induced, could be compared with other learning theories in a variety of ways. Only in the case of learning by example, however, has research progressed sufficiently far to enable the theory to account for what is learned as well as specifying conditions under which learning occurs. For this reason, this learning theory is especially practical. In the subsequent sections, the theory is outlined and its most important implications are considered.

II. HOW CAN LEARNING FROM EXAMPLES OCCUR?

Early work by cognitive psychologists on concept learning identified a variety of types of concepts that could be learned according to the logical structure of the corresponding propositional formulae (cf. Bruner, Goodnow, Austin, 1956; Hunt, 1952). Thus, conjunctive concepts are those for which class membership rules are expressible in terms of conjunctions of two or more criterial attributes. Similarly, relational concepts are those requiring specific relations to hold among objects or features in all examples of the concept. This taxonomy stimulated a vast amount of related research. In particular, the type (logical structure) of concepts, strategies for hypothesis testing, effects of positive or negative

examples, and amount and delay of feedback were heavily investigated.

Recently, a new approach to the learning problem has been taken stemming from combined interests in computer-based learning and cognitive models of human learning. The thrust of this research is to develop mechanisms that can model the kind of learning people exhibit when provided with examples of general rules. This change in emphasis has produced a new orientation to the basic issues. The chief variables of interest in the current research are what kinds of logical descriptions actually underlie cognitive processes and how these can be inferred directly from examples (cf. Hayes-Roth, 1976, 1977(a); Hayes-Roth & Hayes-Roth, 1977; Hayes-Roth & McDermott, 1976(a), 1976(b); Michalski, 1973; Vere, 1975, 1977; Waterman, 1975(a), 1975(b); Williams, 1972; Winston, 1975).

Several researchers have come to the conclusion that a significant amount of human knowledge is representable in two basic conceptual forms. One is a structural representation of a pattern, such as a propositional formula describing the prototypic triangle (see above) or a predicate calculus expression representing a prototypic active sentence that can be transformed into passive voice. The second type of knowledge structure is a production rule that relates two pattern descriptions by a transformational rule: whenever the pattern called the "condition" is matched by events in memory, the

pattern called the 'action' component is executed. Such [condition->action] rules have now been successfully employed to model a wide range of intellectual functions, including syntactic and semantic processing in connected-speech understanding (Hayes-Roth & Mostow, 1976; Hayes-Roth, Mostow, & Fox, 1977), human problem solving (Newell, 1973; Newell & Simon, 1972), and medical diagnosis (Shortliffe, 1976).

The benefit of such an approach to knowledge representation is multiplied considerably by the fact that both of the proposed types of knowledge structures can be induced from examples. Extending and generalizing the early work on conjunctive concept learning, an operational theory has been developed for learning structural representations of patterns and production rules from examples. The essence of the theory is the same as suggested by Galton in 1907 in the form of a "composite photograph" theory of concept formation. Galton converted the contemporary technology of photography into a theory for explaining how it was possible to combine various images of the same object into a generalized form that could serve as a template for recognizing all variants of the basic image. His theory was, simply, that the memory representation of an object's image was like a photographic transparency. Thus, by overlaying multiple representations derived from different views of the same object, a composite photograph would emerge in which the invariant properties would be accentuated and properties that differed among alternative views would be attenuated. The collection of emergent common

properties would constitute a template defining the concept. Subsequent images of the same object would necessarily match this template, because they too would exhibit properties criterial to the general concept.

Of course, Galton's theory was little more than a metaphor. Combining Galton's key insight with contemporary technology, however, yields a powerful model of the learning process (Hayes-Roth, 1976, 1977(a); Hayes-Roth & McDermott, 1976(b)). In brief, the learning process is modeled in the following way. At any moment, the learner is viewed as possessing a current stock of predicates that he/she uses to delimit, label, describe and relate objects, actions, and events. This knowledge is conceived of as a set of properties and case frames applied in forming descriptions of instructional examples. A formal methodology has been developed to convert such case relational descriptions of events into equivalent graph structures. Then, by comparing the graph representations of several illustrations of the same concept or rule and extracting the subgraph that is common to all (or many) of them, an hypothesis about the criterial properties and relations of the general class is generated. As Galton proposed, different examples of the same class concept will all exhibit the criterial properties constituting the general rule. These are directly identified by finding the structural sub-representation that is common to the various descriptions. Thereafter, the learner can recognize novel instances of the general rule by detecting that

the stimulus manifests those properties inferred to be criterial. Furthermore, newly learned patterns (like "triangle") increase the learner's stock of descriptive predicates.

At the same time as this process of learning by comparison identifies the commonalities among different examples of the same rule, one of its by-products is a discovery that particular sets of different properties reliably occur as alternatives in similar descriptive contexts. In fact, categories (such as "noun" and "adjective" are immediately identifiable by noting that specific sets of values occur in comparable contexts in different examples (e.g., "red" or "white" or "blue" all can precede "house" or "flag" or "paint"). Once such categories are identified, the learner can describe subsequent events using these novel properties. Thus, both concept and category learning produce an expansion of the set of predicates available to the learner.

Finally, transformations or problem solving production rules can be induced from before-and-after pairs illustrating their application. In this case, antecedent and consequent components are compared to their respective counterparts in several examples. In this way, general [condition->action] patterns common to all examples are inferred. These learned rules explain how variables are bound and substitutions should be made from constants in a problem context (the condition) to the associated action.

As an illustration, consider the problem of inducing an

unknown rule of transformational grammar from the following three antecedent-consequent example pairs.

(1) "The little man sang a lovely song."-->
'A lovely song was sung by the little man.'

(2) "A girl hugged the motorcycles."-->
"The motorcycles were hugged by a girl."

(3) "People are stopping friendly policemen."-->
"Friendly policemen are being stopped by people."

The relational descriptions of these sentence pairs consist of three types of components: (1) syntactic phrase structures and markers (e.g., NUMBER:SINGULAR, TENSE:PRESENT); (2) a property which distinguishes elements of the antecedent sentence from elements of the consequent sentence (EVENT:e1 and ANTECEDENT:e1 as opposed to EVENT:e2 and CONSEQUENT:e2); and (3) same-type relations joining any pair of antecedent and consequent syntactic components which are identical types (i.e., are distinct tokens of the same type or, equivalently, are the roots of identical directed phrase structure graphs). The case frame description of the first sentence pair is shown below.

```
{{ANTECEDENT:e1, CONSEQUENT:e2},
 {S:s1, NP:np11, VP:vp1, EVENT:e1},
 {S:s2, NP:np21, VP:vp2, EVENT:e2},
 {NP:np11, DET:the1, ADJ:little1, NOUN:noun11, EVENT:e1},
 {NP:np21, DET:a1, ADJ:lovely1, NOUN:noun21, EVENT:e2},
 {NOUN:noun11, NST:man1, NUMBER:n11, EVENT:e1},
 {NOUN:noun21, NST:song1, NUMBER:n12, EVENT:e2},
 {SINGULAR:n11, EVENT:e1},
 {SINGULAR:n12, EVENT:e2},
 {VP:vp1, AUX:aux11, VERB:verb11, NP:np22, EVENT:e1},
 {SAME!NP:np21, SAME!NP:np22},
 {NP:np22, DET:a2, ADJ:lovely2, NOUN:noun22, EVENT:e1},
 {SAME!NOUN:noun21, SAME!NOUN:noun22},
 {NOUN:noun22, NST:song2, NUMBER:n13, EVENT:e1},
```

```
{SINGULAR:n13, EVENT:e1},
{VP:vp2, AUX:aux12, PB:pb1, VERB:vert12, PP:pp1, EVENT:e2},
{AUX:aux11, AUX3T:have1, TENSE:t11, NUMBER:n15, EVENT:e1},
{AUX:aux12, AUX3T:have2, TENSE:t12, NUMBER:n16, EVENT:e2},
{SAME!AUX:aux11, SAME!AUX:aux12},
{VERB:vert11, VST:sing1, TENSE:t21, NUMBER:n15, EVENT:e1},
{VERB:vert12, VST:sing2, TENSE:t22, NUMBER:n16, EVENT:e2},
{SAME!VERB:vert11, SAME!VERB:vert12},
{PB:pb1, PB3T:be1, TENSE:t23, NUMBER:n16, EVENT:e2},
{SAME!TENSE:t11, SAME!TENSE:t12},
{SAME!TENSE:t21, SAME!TENSE:t22, SAME!TENSE:t23},
{SINGULAR:n15, EVENT:e1},
{SINGULAR:n16, EVENT:e2},
{PRESENT:t11, EVENT:e1},
{PRESENT:t12, EVENT:e2},
{PAST-PART:t21, EVENT:e1},
{PAST-PART:t22, PAST-PART:t23, EVENT:e2},
{PP:pp1, PREP:by1, NP:np12, EVENT:e2},
{SAME!NP:np11, SAME!NP:np12},
{NP:np12, DET:the2, ADJ:little2, NOUN:noun12, EVENT:e2},
{SAME!NOUN:noun11, SAME!NOUN:noun12},
{NOUN:noun12, VST:man2, NUMBER:n14, EVENT:e2},
{SAME!NUMBER:n11, SAME!NUMBER:n12, SAME!NUMBER:n13},
{SAME!NUMBER:n14, SAME!NUMBER:n15, SAME!NUMBER:n16},
{SINGULAR:n14, EVENT:e2},
{THE:the1, EVENT:e1},
{THE:the2, EVENT:e2},
{SAME!WORD:the1, SAME!WORD:the2},
{LITTLE:little1, EVENT:e1},
{LITTLE:little2, EVENT:e2},
{SAME!WORD:little1, SAME!WORD:little2},
{MAN:man1, EVENT:e1},
{MAN:man2, EVENT:e2},
{SAME!WORD:man1, SAME!WORD:man2},
{HAVE:have1, EVENT:e1},
{HAVE:have2, EVENT:e2},
{SAME!WORD:have1, SAME!WORD:have2},
{SING:sing1, EVENT:e1},
{SING:sing2, EVENT:e2},
{SAME!WORD:sing1, SAME!WORD:sing2},
{A:a1, EVENT:e1},
{A:a2, EVENT:e2},
{SAME!WORD:a1, SAME!WORD:a2},
{LOVELY:lovely1, EVENT:e1},
{LOVELY:lovely2, EVENT:e2},
{SAME!WORD:lovely1, SAME!WORD:lovely2},
{SONG:song1, EVENT:e1},
{SONG:song2, EVENT:e2},
{SAME!WORD:song1, SAME!WORD:song2},
{BE:be1, EVENT:e2},
{BY:by1, EVENT:e2}}
```

This corresponds to Figure 1. When the descriptions of all three sentence pairs are compared, the learning algorithm induces the rule illustrated in Figure 2. Here, same-type relations that were common to all examples have been retained and are represented by arrows linking subgraphs in the antecedent sentence structure with corresponding subgraphs in the consequent sentence structure. These correspondences represent identical quantified variables in the left and right-hand sides of the inferred production. When the rule is applied, the actual parameter in a stimulus event which matches the antecedent will be bound to the corresponding quantified variable and should be substituted into the corresponding locus of the consequent structure. If this is done, the inferred production will effect the active-to-passive transformation rule.

This theory has led to development of several successful computer learning algorithms. These have been used to infer concepts in a variety of language processing tasks, including: learning to discriminate among vowels occurring in natural, connected speech (Burge & Hayes-Roth, 1976); learning to recognize words by discovering finite-state network representations comprising sequences of sets of alternative, acoustically labeled, phones evident in examples (Fox & Hayes-Roth, 1976; Hayes-Roth & Burge, 1976); induction of general syntactic rules for active-to-passive, equi-noun phrase deletion, and other transformations (Hayes-Roth, 1977(a); Hayes-Roth & McDermott, 1976(b)); and induction of grammatical and semantic

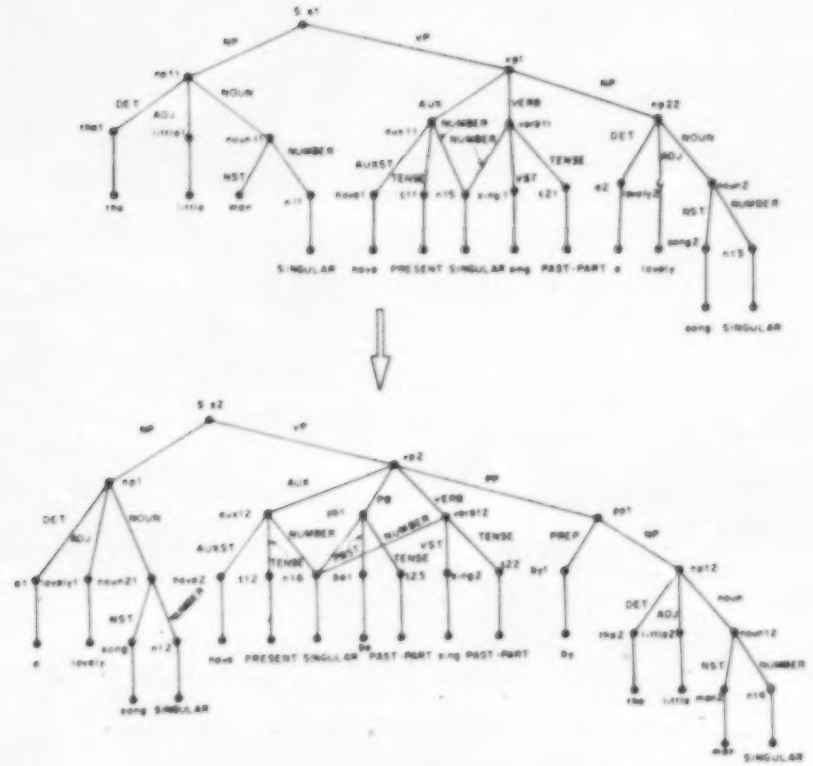

Figure 1. Graphical representation of the first example of the active-to-passive transformational rule.

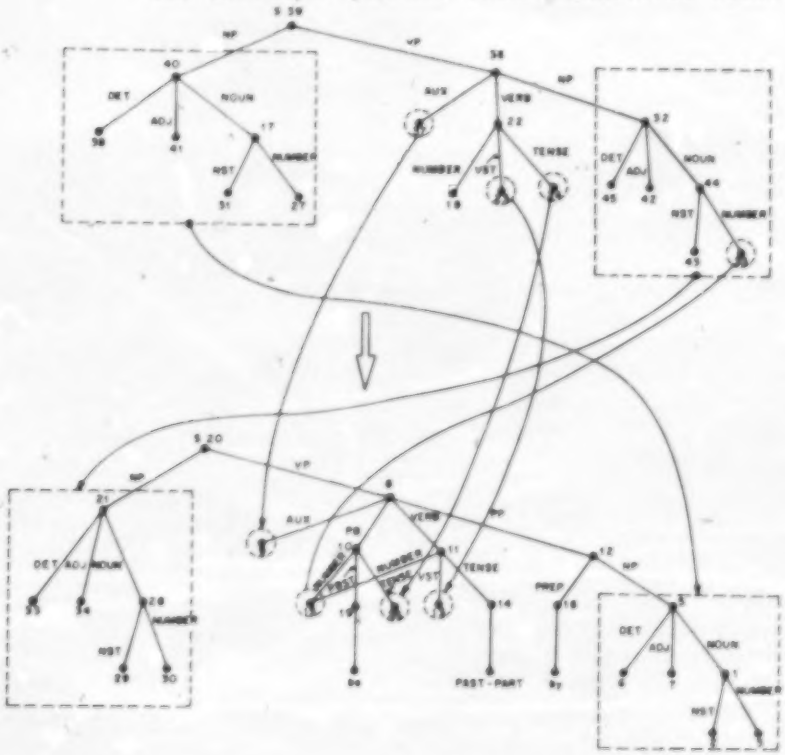

Figure 2. Graphical representation of the induced active-to-passive transformational rule.

categories of words and phrases from a corpus of language (Hayes-Roth, 1976).

Application of the theory to human learning tasks has shown that concept learning occurs when criterial combinations of properties common to examples of the same class are encoded and strengthened in memory (Hayes-Roth & Hayes-Roth, 1976, 1977(a)). Human recognition and classification performance is most accurately predicted by postulating that decisions are based on the memory strengths and cue validities of combinations of properties occurring in earlier examples. The significance of property combinations abstracted from instructional examples is a persistent and powerful phenomenon. Even when subjects are taught to apply a general rule and are subsequently trained by practicing with specific examples, their recollection of training instances and classification of subsequent items are more accurately predicted by the proposed learning theory than by theories postulating storage of an abstract or symbolic prototype plus transformational rules that accommodate variations from prototypes (Hayes-Roth & Hayes-Roth, 1977).

III. WHEN DOES LEARNING BY EXAMPLE OCCUR?

The results of the research mentioned above suggest that we learn general rules, in the form of criterial combinations of properties and relations among objects, whenever we practice

applying a rule or try to learn an unknown rule with instructive examples. This learning apparently occurs even when ostensibly unnecessary; i.e., if a general rule for a new concept is presented first, subsequent practice with specific examples of the rule results in learning of generalized representations that are abstracted directly from observed properties and relations of the specific examples. In short, if examples are used, they have a significant effect on what is learned (Hayes-Roth & Hayes-Roth, 1976, 1977).

IV. HOW CAN POSITIVE EFFECTS OF LEARNING BE PROMOTED?

This theory of learning predicts that each instructional example causes learning of relevant and irrelevant criteria that are consistent with it. Thus, undesirable learning must be controlled. Two methods for limiting learning of irrelevant generalities are: (1) present examples with the fewest irrelevant characteristics; and (2) insure that multiple examples of the same general rule vary maximally on irrelevant dimensions.

The theory also predicts that examples are the most powerful mechanism for instilling effective rule-directed behavior. They apparently influence behavior more than presentation of formal statements of general rules. Thus, learning can be most facilitated by first identifying the criterial properties and relationships of the knowledge to be acquired and then insuring that each example manifests all criterial characteristics.

Other conditions that might promote learning by example are discussed in the next section.

V. RESEARCH DIRECTIONS AND POTENTIAL APPLICATIONS

First, consider the important questions that are raised by these results. Among these are the following:

> Is teaching with counterexamples effective? Results of earlier concept learning studies suggest that subjects have difficulty learning from negative information. This would be predicted by the current theory too.
>
> What presentation sequence of rules and examples most promotes rapid learning?
>
> How can examples be constructed to facilitate abstraction of more than one principle at a time?
>
> Are informationally rich, embellished examples more effective or less effective than impoverished, bare-bones examples constructed to minimize irrelevant information?
>
> Are variants of the teaching by example paradigm, such as teaching by analogy or anecdote, effective? How can the positive or negative effects associated with these alternatives be predicted?

Finally, consider the implications of this work for instructional applications. The following instructional paradigm appears to be warranted: first, the pattern characteristic or [condition->action] rule to be learned should be analyzed to identify its criterial properties and relations; competence in detecting and coding these attributes should then be established; multiple and varied illustrations should be presented that manifest, in a salient way, the criterial properties and relations; and the multiple examples should be explicitly

compared to highlight their commonalities. On the basis of evidence accumulated to date, this paradigm can be expected to promote rapid and effective learning.

REFERENCES

Bruner, J. S., Goodnow, J. J., and Austin, G. A. A study of thinking. New York: Wiley, 1956.

Burge, J., & Hayes-Roth, F. A novel pattern learning and classification procedure applied to the learning of vowels. Proceedings of the 1976 I.E.E.E. International Conference on Acoustics, Speech and Signal Processing. Philadelphia, 1976, 154-157.

Fox, M. S., & Hayes-Roth, F. Approximation techniques for the learning of sequential symbolic patterns. Proceedings of the Third International Joint Conference on Pattern Recognition, Coronado, California, 1976, 616-620.

Galton, F. Inquiries into human faculty and its development. London: Dent, 1907.

Hayes-Roth, B., & Hayes-Roth, F. Plasticity in memorial networks. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 506-522.

Hayes-Roth, B., & Hayes-Roth, F. A comparative psychological investigation of pattern classification theories. Proceedings of the Third International Joint Conference on Pattern Recognition, Coronado, California, 1976, 224-227.

Hayes-Roth, B., & Hayes-Roth, F. Concept learning and the classification and recognition of exemplars. Journal of Verbal Learning and Verbal Behavior, 1977 (in press).

Hayes-Roth, F. A structural approach to pattern learning and the acquisition of classificatory power. Proceedings of the First International Joint Conference on Pattern Recognition, New York, 1973, 343-355.

Hayes-Roth, F. Schematic classification problems and their solution. Pattern Recognition, 1974, 6, 105-113.

Hayes-Roth, F. Patterns of induction and related knowledge acquisition algorithms. In C. Chen (Ed.), Pattern recognition and artificial intelligence. New York: Academic Press, 1976.

Hayes-Roth, F. Uniform representations of structured patterns and an algorithm for the induction of contingency-response rules. Information and Control, 1977(a), 33, 87-116.

Hayes-Roth, F. The role of partial and best matches in knowledge systems. In D.A. Waterman & F. Hayes-Roth (Eds.), Pattern-directed inference systems. New York: Academic Press, 1977(t), in press.

Hayes-Roth, F., & Burge, J. Characterizing syllables as sequences of machine-generated labelled segments of connected speech: a study in syntolic pattern learning using a conjunctive feature learning and classification system. Proceedings of the Third International Joint Conference on Pattern Recognition, Coronado, California, 1976, 431-435.

Hayes-Roth, F., & McDermott, J. Learning structured patterns from examples. Proceedings of the Third International Joint Conference on Pattern Recognition, Coronado, California, 1976(a), 419-423.

Hayes-Roth, F., & McDermott, J. Knowledge acquisition from structural descriptions. Pittsburgh: Working paper, Department of Computer Science, Carnegie-Mellon University, 1976(t).

Hayes-Roth, F., & Mostow, D. J. Syntax and semantics in a distributed logic speech understanding system. Proceedings of the 1976 I.E.E.E. International Conference on Acoustics, Speech and Signal Processing, Philadelphia, 1976, 421-424.

Hayes-Roth, F., Mostow, D. J. & Fox, M. Understanding speech in the Hearsay-II system. In L. Bolc (Ed.), Natural language communication with computers. Berlin: Springer-Verlag, 1977.

Hunt, E. B. Concept learning: an information processing approach. New York: Wiley, 1962.

Michalski, R. S. AQVAL/1--Computer implementation of a variable valued logic system VL1 and examples of its application to pattern recognition. Proceedings of the First International Joint Conference on Pattern Recognition, Washington, D.C., 1973, 3-17.

Newell, A. Production systems: models of control structures. In W. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Newell, A., & Simon, H. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Plotkin, G. D. A note on inductive generalization. In B. Meltzer and D. Michie (Eds.), Machine Intelligence, vol. 5. New York: American Elsevier, 1970.

Plotkin, G. D. A further note on inductive generalization. In B. Meltzer and D. Michie (Eds.), Machine Intelligence, vol. 6. New York: American Elsevier, 1971.

Shortliffe, T. MYCIN: A program for computer-based medical consultations. New York: Elsevier North-Holland, 1976.

Vere, S. A. Induction of concepts in the predicate calculus. Proceedings of the Fourth International Joint Conference on Artificial Intelligence, Tbilisi, U.S.S.R., 1975, 281-287.

Vere, S.A. Induction of relational productions. In D.A. Waterman and F. Hayes-Roth (Eds.), Pattern-directed inference systems. New York: Academic Press, 1977 (in press).

Waterman, D. A. Adaptive production systems. Proceedings of the Fourth International Joint Conference on Artificial Intelligence, Tbilisi, U.S.S.R., 1975(a), 296-303.

Waterman, D. A. Serial pattern acquisition: a production system approach. Pittsburgh: CIP Working Paper, Department of Psychology, Carnegie-Mellon University, 1975(b).

Williams, D. S. Computer program organization induced from problem examples. In H. A. Simon & L. Siklossy (Eds.), Representation and meaning. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Winston, P. H. Learning structural descriptions from examples. In P. H. Winston (Ed.), Psychology of computer vision. New York: McGraw-Hill, 1975.